DOI:10.11931/guihaia.gxzw201907022

小立碗藓扩展蛋白基因家族的鉴定与生物信息学分析

蓝雨纯[1]，黄彬[1]，韦娇[1]，姜山[2*]

（1. 贵州师范大学 生命科学学院，贵阳 550001；2. 贵州师范大学 国际教育学院，贵阳 550001）

摘要：扩展蛋白（Expansins, EXP）是一类基因家族，几乎参与了植物发育的全过程，从种子萌发到果实成熟都有扩展蛋白的参与。利用生物信息学的方法对小立碗藓（*Physcomitrella patens*）Expansin 基因家族成员进行鉴定，并分析其基因结构、染色体定位及系统发生关系；结果表明小立碗藓基因组中含有 Expansin A（EXPA）32 个、Expansin-like A(EXLA)6 个，并未发现 Expansin-like B（EXLB）及 Expansin B（EXPB）。扩展蛋白氨基酸序列长度在 228 ~ 290 之间，编码蛋白质具有两个保守的结构域 Pollen_allerg_1 和 DPBB_1。蛋白质亚细胞定位预测结果表明：运用 CELLO 在线工具预测发现小立碗藓中约 4/5 的 EXP 家族基因定位于细胞外；而 Euk-mPLoc 预测结果则显示小立碗藓 EXP 基因家族成员全定位于细胞外。基因结构分析表明，小立碗藓中约 68% Expansin 基因有含有 1 ~ 3 个内含子。本研究分析了小立碗藓扩展蛋白基因家族的基本信息，可为深入研究其扩展蛋白基因的分子进化与生物学功能奠定基础。

关键词：小立碗藓，扩展蛋白，基因家族，生物信息学

中图分类号：Q943　　　　　　　　　　　**文献标识码：**A

Identification and bioinformatic analysis of the expansin gene family in *Physcomitrella patens*

LAN Yuchun[1], HUANG Bing1, WEI Jiao[1], JIANG Shan[2*]

（1. *School of Life Sciences*, *Guizhou Normal University*, Guiyang 550001, China; 2. *School of International Education*, *Guizhou Normal University*, Guiyang 550001, China）

Abstrac: Expansins (EXP) is a family of genes that are involved in the whole process of plant development. From seed germination to fruit ripening, there are Expansin involved. Bioinformatics methods were used to identify the members of the Expansin gene family of *Physcomitrella patens (P. patens),* and their gene structure, chromosomal location and phylogenetic relationship were analyzed. The results indicated that the genome of the *P. patens* contains 32 Expansin A (EXPA) and 6 Expansin-like A (EXLA). Expansin-like B (EXLB) and Expansin B (EXPB) were not found. The Expansin amino acid sequence is between 228 aa and 290 aa in size, and the encoded protein has two conserved domains, Pollen_allerg_1 and DPBB_1. The results of protein subcellular localization prediction showed that about 4/5 of the EXP family genes in the *P. patens* were located outside the cell using the CELLO online tool;while the Euk-mPLoc prediction showed that the EXP gene family member was fully localized in extracellular. Gene structure analysis showed that about 68% of the Expansin gene in *P. patens* contains 1-3 introns.

基金项目：国家自然科学基金（0515049）[Supported by the National Natural Science Foundation of China (0515049）]。

作者简介：蓝雨纯(1995 -)，女，四川德阳中江县人，硕士研究生，主要从事植物分子生物学研究，(E-mail）903417069@qq.com。

***通信作者：**姜山，博士，教授，研究方向为植物病理学，(E-mail）kyosan200312@hotmail.com。

This study analyzed the basic information of the *P. patens* extended protein gene family, which lays a foundation for further study of the molecular evolution and biological functions of its extended protein genes.
Key words: *Physcomitrella patens*, Expansin, gene family, bioinformatics

植物细胞壁是植物细胞中是一种极其重要的结构。细胞壁能够决定细胞的形状，为细胞提供基本的机械支撑和硬度，是细胞抗击病原体的最后一道屏障(Cosgrove, 2005)。扩展蛋白，即植物细胞壁松弛蛋白(Expansins, EXP)是一类细胞壁蛋白，主要存在于各植物细胞组织中,通过引起细胞壁组分间的松驰和细胞壁柔韧性增加，参与植物的许多生长发育过程。

在对黄瓜(*Cucumic sativus*)下胚轴细胞壁酸诱导伸展的研究实验中，Cosgrove et al. (2002)首次发现 EXP。实验结果显示：EXP 可以使经热钝化后的细胞壁恢复伸展活性，但不具有溶菌酶活性，并且达到几乎和离体细胞壁本身伸展活性相当的程度，以此推测其具有调节植物细胞壁的伸展性的功能（Javier et al., 1992）。

实验证实 EXP 与许多生长发育过程相关，例如：Cosgrove et al. (2002)实验发现果实成熟软化和花粉管延伸有关；Belfield et al. (2005)研究表明 EXP 也与叶片脱落有关；除此之外， Han et al. (2004)研究也显示 EXP 基因在植物抗逆中也有重要作用。早期研究表明，EXP 以酶催化的作用方式，使细胞壁组分间疏松，细胞伸展，增强细胞柔韧性，以此缓解细胞在不良环境下的压力(Cosgrove, et al., 1997)。

近年来，随着基因组测序技术的普及，越来越多的扩展蛋白在其它物种中被鉴定。依据进化关系，可将扩展蛋白基因家族划分为四个亚家族 α、β、γ 和 δ。如今，科学研究在多种植物的不同生长阶段都发现扩展蛋白存在。α-和 β-expansin 家族分别被重命名为 EXPA 与 EXPB，γ-expansin，称为 expansin-like A（EXLA）和 δ-expansin 称为 expansin-like B（EXLB）(金慧清等，2006)。EXPA 和 EXPB 亚家族蛋白具有细胞壁伸展活性，参与细胞壁伸长和其它发育过(Hans, et al., 2004)；但 EXLA 和 EXLB 两类亚家族仅仅知道其基因序列，暂无相关实验证明 EXLA 和 EXLB 两类亚家族蛋白具有增强细胞壁韧性的功能(Sampedro, et al., 2006)。

小立碗藓(*Physcomitrella patens*)是非维管束植物类群的模式植物，进化地位特殊，全基因组也已测序完成，是研究功能基因组学、发育生物学、植物生理、系统进化理想材料。小立碗藓基因组大小为 511 Mb，共含 27 条染色体(Rensing, et al., 2002)，有着易与外源基因发生同源重组，易培养且生长周期较短，表型易观察等优势。本研究基于 2007 年小立碗藓的全基因组序列测序工作完成(http: //www.cosmoss.org/)(Rensing, et al., 2008)，利用生物信息学的方法筛选鉴定小立碗藓扩展蛋白基因家族成员，并对其进行进化树构建，染色体定位、基因结构分析等。为后续研究该基因家族提供了参照。

1 材料和方法

1.1 实验材料

本研究以模式植物小立碗藓为研究对象，小立碗藓全基因组数据下载于 Ensembl Plants 数据库(Bolser, et al., 2017)（http://plants.ensembl.org/index.html）中最新数据。

1.2 小立碗藓 Expansin 基因家族的鉴定

由 Ensembl 数据库（http://asia.ensembl.org/index.html）中获得小立碗藓基因组相关数据；从 Pfam 数据库（http://pfam.xfam.org/）下载蛋白保守结构域的隐马科夫模型文件(*.hmm)。利用 HMMER 软件(Finn, et al., 2011)中 hmmsearch 工具搜索小立碗藓全基因组预测蛋白序列中含有 Pollen_allerg_1(PF01357)和

DPBB_1(PF03330)保守结构域的序列(李昊阳等，2014)，筛选 E-value < 1 ×10^{-20} 的序列作为初筛序列。利用初筛序列及 HMMER 软件重新构建小立碗藓的特异性 hmm 文件，并利用该 hmm 文件再次进行筛选(E-value < 0.01)，筛选所得序列为备用序列。为保证搜索结果的准确性，通过 SMART 网站（http://smart.embl-heidelberg.de/）对备用蛋白进行手动筛选(Letunic, et al., 2012)，以手动剔除不含 Pollen_allerg_1 和 DPBB_1 两个保守结构域的序列。

1.3 小立碗藓扩展蛋白基因家族的蛋白特征分析及亚细胞定位

将鉴定得到的小立碗藓扩展蛋白家族蛋白序列，用过 NCBI（https://www.ncbi.nlm.nih.gov/）数据库查找其分子量(Mw)、等电点(PI)等数据。通过 Cello 在线工具（http://cello.life.nctu.edu.tw/）及 Euk-mPLoc 在线网站（http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/）预测小立碗藓 EXP 基因家族的亚细胞定位。

1.4 小立碗藓 Expansin 家族系统进化树的构建

利用小立碗藓 Expansin 基因的蛋白序列，与 TAIR 据库（https://www.arabidopsis.org/）及 NCBI 数据库（https://www.ncbi.nlm.nih.gov/）中搜索到的拟南芥(*Arabidopsis thaliana*)中 Expansin 基因蛋白序列，应用多序列比对工具 ClustalX 构建系统进化树，以氨基酸全序列联配的结果为基础，用 MEGA 7.0 程序生成。采用程序 Maximum Likelihood 法，校验参数为 Bootstrap = 1 000。

1.5 小立碗藓 Expansin 基因保守基序及基因结构分析

通过 MEME 网站检测，小立碗藓 Expansin 家族基因中所存在相似度较高的基序（motif）；并利用基因组注释文件，获得小立碗藓 Expansin 基因家族的内含子及外显子分布情况。最后利用 TBtools 软件对其进行可视化分析。

1.6 小立碗藓 Expansin 基因的染色体位置图

以 Phytozome 数据库（https://phytozome.jgi.doe.gov/pz/portal.html）提供的相关基因信息，确定 Expansin 基因的染色体定位。然后通过 MapInspect 软件作图(易吉明等，2015)。

2 结果与分析

2.1 小立碗藓 Expansin 基因家族成员信息

通过 HMMER 软件对小立碗藓全基因组预测蛋白序列进行搜索，共获得 91 个备用蛋白。运用 SMART 网站手动鉴定备用蛋白后，最终获得 38 个小立碗藓扩展蛋白（表 1）。与双子叶模式植物拟南芥中鉴定的扩展蛋白数量持平(Seader, et al., 2016)。

通过蛋白特征分析，小立碗藓 38 个 Expansin 基因的分子量为 25.14 kDa 到 73.95 kDa，其中约 87%的扩展蛋白分子量处于 20~30 kDa，等电点最小为 4.14。这些小立碗藓 Expansin 基因编码的蛋白包含 228 ~290 个氨基酸。亚细胞定位预测结果显示：运用 CELLO 在线工具预测发现小立碗藓中约 1/5 的 EXP 家族基因定位于细胞周质；而 Euk-mPLoc 预测结果则显示小立碗藓 EXP 基因家族全定位于细胞外。Pfam 数据库(http://pfam.xfam.org)验证结果显示 38 个扩展蛋白同时具有 Pollen_allerg_1 和 DPBB_1 特征结构域。

表 1 小立碗藓 Expansin 基因家族成员基本信息
Table 1 Identification of Expansin gene family members from *Physcomitrella patens*

	亚细胞定位

基因名称 Gene name	基因号 Gene ID	染色体置 Location	等电点 PI	分子量 Mw (kDa)	E-valu值 E-value	氨基酸长度 Length of protein (aa)	内含子数目 Intron	mRNA长度 mRNA length	Subcellular localization CELLO 预测 CELLO prediction	Subcellular localization Euk-mPLoc 预测 Euk-mPLoc prediction
PpEXPA1	GeneID: 112285422	Chr8:8694013--8696181	11.53	23.51	2.00E-122	269	1	1 496	细胞外 Extracellular（3.123*）	细　胞　外 Extracellular
PpEXPA2	GeneID: 112290281	Chr13:14311266--14313155	8.36	26.29	1.00E-130	262	1	1 366	细胞外 Extracellular（3.522*）	细　胞　外 Extracellular
PpEXPA3	GeneID: 112289447	Chr12:3275318--3277508	8.63	26.35	4.00E-133	255	2	1 272	细胞外 Extracellular（3.998*）	细　胞　外 Extracellular
PpEXPA4	GeneID: 112288396	Chr11:7792124--7794136	8.1	26.74	8.00E-150	262	1	1 516	细胞外 Extracellular（2.295*）	细　胞　外 Extracellular
PpEXPA5	GeneID: 112273392	Chr20:3549449--3552027	6.03	26.79	1.00E-132	286	3	1 768	细胞外 Extracellular（3.334*）	细　胞　外 Extracellular
PpEXPA6	GeneID: 112294930	Chr18:13986071--13988541	10.05	27.14	5.00E-147	253	2	1 902	细胞外 Extracellular（4.194*）	细　胞　外 Extracellular
PpEXPA7	GeneID: 112276452	Chr24:10147408--10149750	9.25	27.37	2.00E-133	272	3	1 738	细胞外 Extracellular（3.342*）	细　胞　外 Extracellular
PpEXPA8	GeneID: 112285449	Chr8:9701908--9704893	6.7	27.39	8.00E-127	252	1	2 193	细胞外 Extracellular（3.221*）	细　胞　外 Extracellular
PpEXPA9	GeneID: 112290417	Chr13:5618875--5622902	7.49	27.43	0	263	2	2 442	细胞外 Extracellular（2.979*）	细　胞　外 Extracellular
PpEXPA10	GeneID: 112291187	Chr14:11969561--11970412	7.49	27.43	8.00E-158	261	0	1 337	细胞周质 Periplasm (2.396*)	细　胞　外 Extracellular
PpEXPA11	GeneID: 112274102	Chr21:1212542--1214849	9.16	27.58	5.00E-167	253	2	1 697	细胞外 Extracellular（4.075*）	细　胞　外 Extracellular
PpEXPA12	GeneID: 112285508	Chr8:678047--680168	8.1	27.71	4.00E-133	257	2	1 325	细胞外 Extracellular（3.057*）	细　胞　外 Extracellular

PpEXPA13	GeneID: 112284709	Chr7:8846261--8849088	9.01	27.82	0	274	0	1 598	细胞周质 Periplasm (2.983*)	细胞外 Extracellular
PpEXPA14	GeneID: 112274755	Chr22:566463--569299	9.83	28.21	0	268	2	2 032	细胞外 Extracellular（3.565*）	细胞外 Extracellular
PpEXPA15	GeneID: 112285508	Chr8:678029--680162	9.83	28.21	4.00E-133	257	2	1 325	细胞外 Extracellular（3.057*）	细胞外 Extracellular
PpEXPA16	GeneID: 112285133	Chr7:8674342--8676137	9.83	28.22	3.00E-114	275	2	1 096	细胞周质 Periplasm (1.761*)	细胞外 Extracellular
PpEXPA17	GeneID: 112294922	Chr18:13965591--13968229	8.8	28.32	1.00E-167	256	2	1 665	细胞外 Extracellular（3.050*）	细胞外 Extracellular
PpEXPA18	GeneID: 11229436	Chr14:12352288--12353576	9.8	28.36	0	261	0	2 689	细胞外 Extracellular（2.716*）	细胞外 Extracellular
PpEXPA19	GeneID: 112274474	Chr21:1250662--1254179	9.47	28.42	0	251	2	1 928	细胞外 Extracellular（3.816*）	细胞外 Extracellular
PpEXP20	GeneID: 112291069	Chr14:12439216--12440893	9.75	28.48	1.00E-180	261	0	1 318	细胞外 Extracellular（2.588*）	细胞外 Extracellular
PpEXPA21	GeneID: 112291079	Chr14:12252757--12254538	5.11	28.67	0	261	1	2 107	细胞外 Extracellular（2.802*）	细胞外 Extracellular
PpEXPA22	GeneID: 112274056	Chr21:1231559--1233452	5.18	29	0	259	2	1 263	细胞外 Extracellular（3.704*）	细胞外 Extracellular
PpEXPA23	GeneID: 112285402	Chr8:8699690--8702014	9.76	29.51	0	290	2	1 496	细胞外 Extracellular（3.147*）	细胞外 Extracellular
PpEXPA24	GeneID: 112295049	Chr18:13972299--13977304	8.96	29.63	1.00E-143	261	2	1 658	细胞外 Extracellular（3.591*）	细胞外 Extracellular
PpEXP25	GeneID: 112279484	Chr3:24972100--24974078	9.5	29.91	1.00E-134	258	2	1 369	细胞外 Extracellular（4.019*）	细胞外 Extracellular
PpEXPA26	GeneID: 112280081	Chr3:11436543--11438886	9.68	29.94	0	255	1	1 634	细胞外 Extracellular（3.010*）	细胞外 Extracellular
PpEXPA27	GeneID: 1122911	Chr14:11965551--11	10.18	30.15	0	261	0	1 337	细胞周质 Periplasm	细胞外 Extracellular

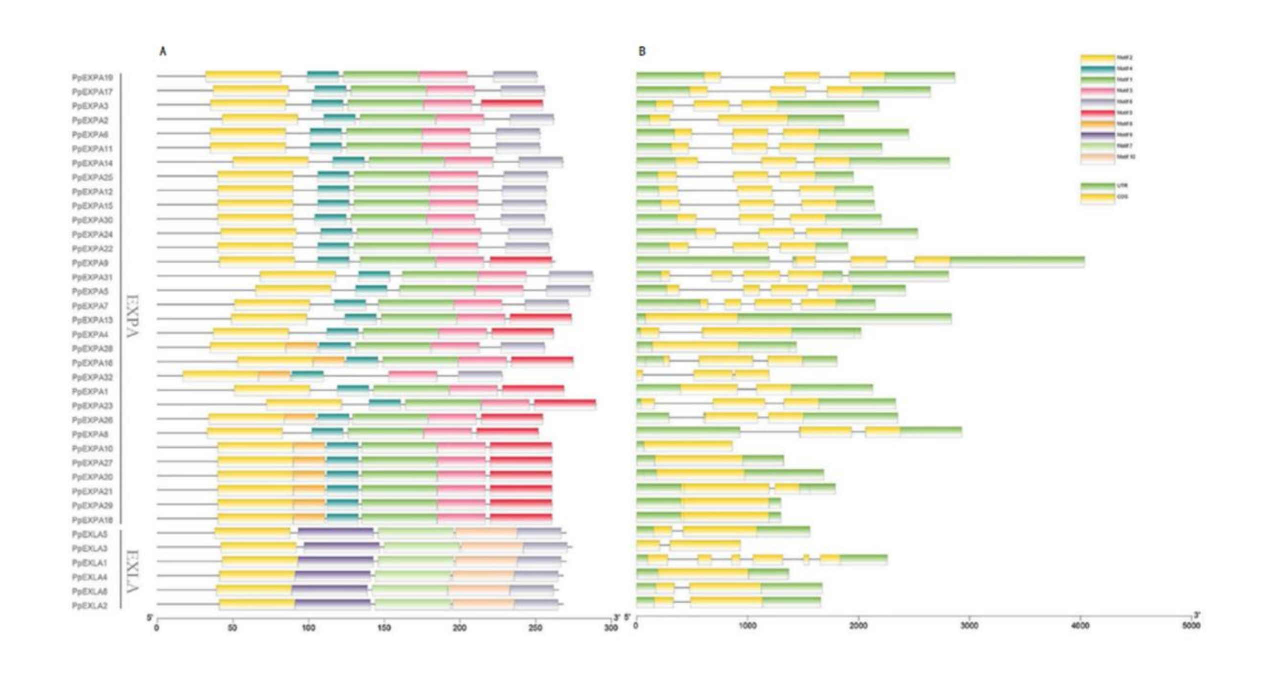
	76	966867							(2.396*)	
PpEXPA28	GeneID: 112284459	Chr7:8724648--8726078	9.27	30.2	5.00E-137	256	0	1 360	细胞周质 Periplasm (2.831*)	细胞外 Extracellular
PpEXPA29	GeneID: 112291122	Chr14:12262439--12263727	6.48	30.69	3.00E-175	261	0	2 107	细胞外 Extracellular（2.716*）	细胞外 Extracellular
PpEXPA30	GeneID: 112274878	Chr22:2293551--2295747	9.24	31.34	6.00E-121	256	2	1 505	细胞外 Extracellular（3.101*）	细胞外 Extracellular
PpEXPA31	GeneID: 112285490	Chr8:10007663-10010516	9.91	31.54	0	288	3	1 678	细胞外 Extracellular（3.876*）	细胞外 Extracellular
PpEXPA32	CDD:215097	Chr7:17045074--17046258	4.14	73.95	2.00E-153	228	2	1 242	细胞外 Extracellular（2.631*）	细胞外 Extracellular
PpEXLA1	GeneID: 112274852	Chr22:11412706--11414985	10.18	25.14	0	270	5	1 342	细胞外 Extracellular（1.864*）	细胞外 Extracellular
PpEXLA2	GeneID: 112274829	Chr22:474094--475773	6.93	27	3.00E-130	268	1	1 347	细胞外 Extracellular（2.385*）	细胞外 Extracellular
PpEXLA3	CDD:215542	Chr2:2168090--2169432	9.87	27.68	7.00E-161	274	1	1 509	细胞周质 Periplasm (1.719*)	细胞外 Extracellular
PpEXLA4	GeneID: 112291165	Chr17:9182556--9183918	5.33	27.85	3.00E-124	268	0	1 369	细胞周质 Periplasm (2.110*)	细胞外 Extracellular
PpEXLA5	GeneID: 112276760	Chr1:26361818--26363694	9.83	28.32	0	270	1	1 441	细胞外 Extracellular（2.326*）	细胞外 Extracellular
PpEXLA6	GeneID: 112294328	Chr17:9178519--9180264	4.91	29.05	1.00E-130	265	1	1 520	细胞周质 Periplasm (2.445*）	细胞外 Extracellular

2.2 小立碗藓 Expansin 基因保守结构域与基因结构分析

通过 MEME 分析小立碗藓 Expansin 基因序列，获得 10 个保守性较高的 motif（图 1：A）。同时，motif 之间的位置具有重要规律，其中发现小立碗藓 EXPA 亚家族除 *PpEXPA32* 出现 motif 1 缺失外，该亚家族其余基因具有 motif 2 -（ motif 8） - motif 4 - motif 1 -motif 3 的相对稳定结构；而 EXLA 亚家族则具有 motif 2 - motif 9 - motif 7 - motif 10 -motif 6 的稳定结构，并未出现 motif 缺失、增加或替换。

分析 38 个小立碗藓 EXP 基因的内含子、外显子结构（图 1：B），发现 EXPA 中约 49%的基因含有 2 个内含子、约 1/5 的基因仅含有 1 个内含子、9%的基因含有 3 个内含子、更有约 22%的 EXPA 亚家族基因

不含内含子。在分析小立碗藓 EXLA 亚家族时发现，EXLA 几乎仅含 1 个内含子，除 *PpEXLA1* 和 *PpEXLA4* 外。其中 *PpEXLA1* 含有 5 个内含子，为 Expansin 基因家族中含有内含子数目最多的基因。



注：A. Expansin 基因的保守基序(motif)分布。10 个 motif 用不同颜色的方框表示(详见表 2)。b.Expansin 基因内含子、外显子分布。绿色方框表示外显子，黑色线条表示内含子。黄色方框表示的是 Expansin 基因的上游/下游区域。外显子的长度可以通过底部的比例尺推断出来。

Note: A. Distributions of conserved motifs in Expansin genes. Ten putative motifs are indicated in different colored boxes (See Table 2 for details.). B. Exon/intron organization of Expansin genes. Green boxes represent exons and black lines with same length represent introns. The upstream/downstream region of Expansin genes are indicated in yellow boxes. The length of exons can be inferred by the scale at the bottom.

图 1 小立碗藓 Expansin 基因保守基序及基因结构分析

Fig. 1 Gene structure and conserved motif analysis of *Physcomitrella patens* Expansin

表 2 Expansin 蛋白的预测 motif 列表

Table 2 List of the putative motifs of Expansin proteins

基序 Motif	E-value 值 E-value	长度 Width	Logo
motif1	2.2e-911	41	
motif2	2.8e-1148	50	

motif3	6.3e-610	34	
motif4	1.1e-521	29	
motif5	6.30E-298	15	
motif6	2.9e-399	21	
motif7	8.40E-146	50	
motif8	2.60E-133	21	
motif9	1.10E-121	49	
motif10	9.20E-107	50	

2.3 小立碗藓扩展蛋白基因家族的系统进化分析

对小立碗藓 Expansin 基因家族进行了氨基酸多序列比对和系统树构建（图 2），分析表明，小立碗藓的 Expansin 基因家族系统进化树有 3 个非常明显的分枝，且小立碗藓 Expansin 基因家族的两个亚基因家族都出现了独立的进化分支。其中 EXPA 亚家族包含两进化分支，说明 EXPA 亚家族的基因在长期的进化过程中，出现了不同的进化方式。另外，小立碗藓中有些 Expansin 基因分支较长，说明这些基因在很早的时候就发生了分化，相应基因序列也已经发生较大的分化，但仍然可以肯定它们之间具有一定的演化关系。

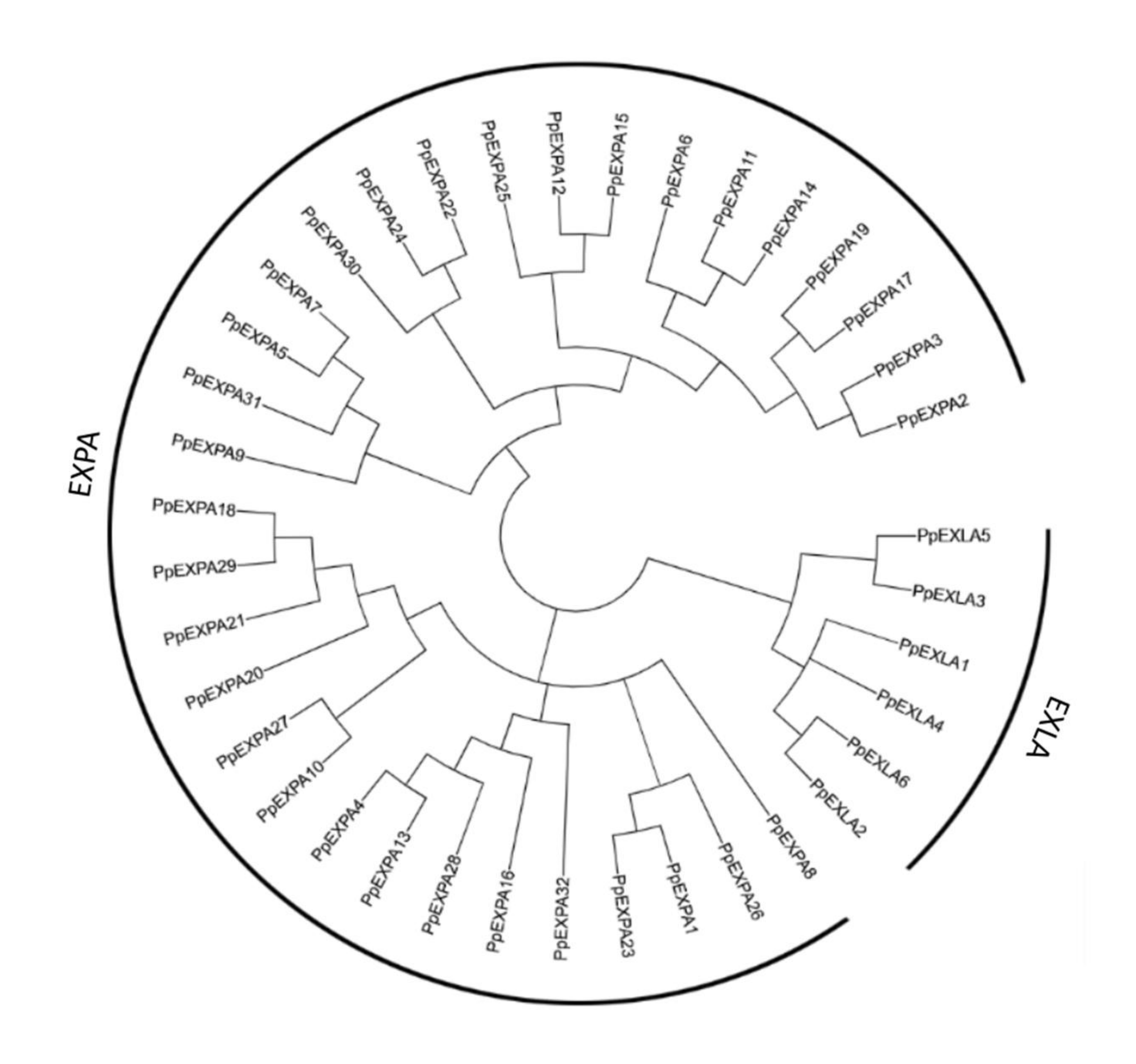


图 2 小立碗藓 Expansin 基因系统进化分析

Fig. 2 Phylogenetic tree of Expansin gene in *Physcomitrella patens*

为研究模式植物小立碗藓与拟南芥基因组中 EXP 之间的进化关系，利用 38 个小立碗藓 EXP 基因和 38 个拟南芥扩展蛋白的氨基酸序列构建了系统发育树(图 3)。从系统发育树可知，小立碗藓 38 个 EXP 基因被明显分为 2 个亚家族。共鉴定种内的直系同源蛋白 12 对，其中有 4 对直系同源蛋白间的 Bootstrap 值为 99，而物种内的旁系同源蛋白数目为 1。

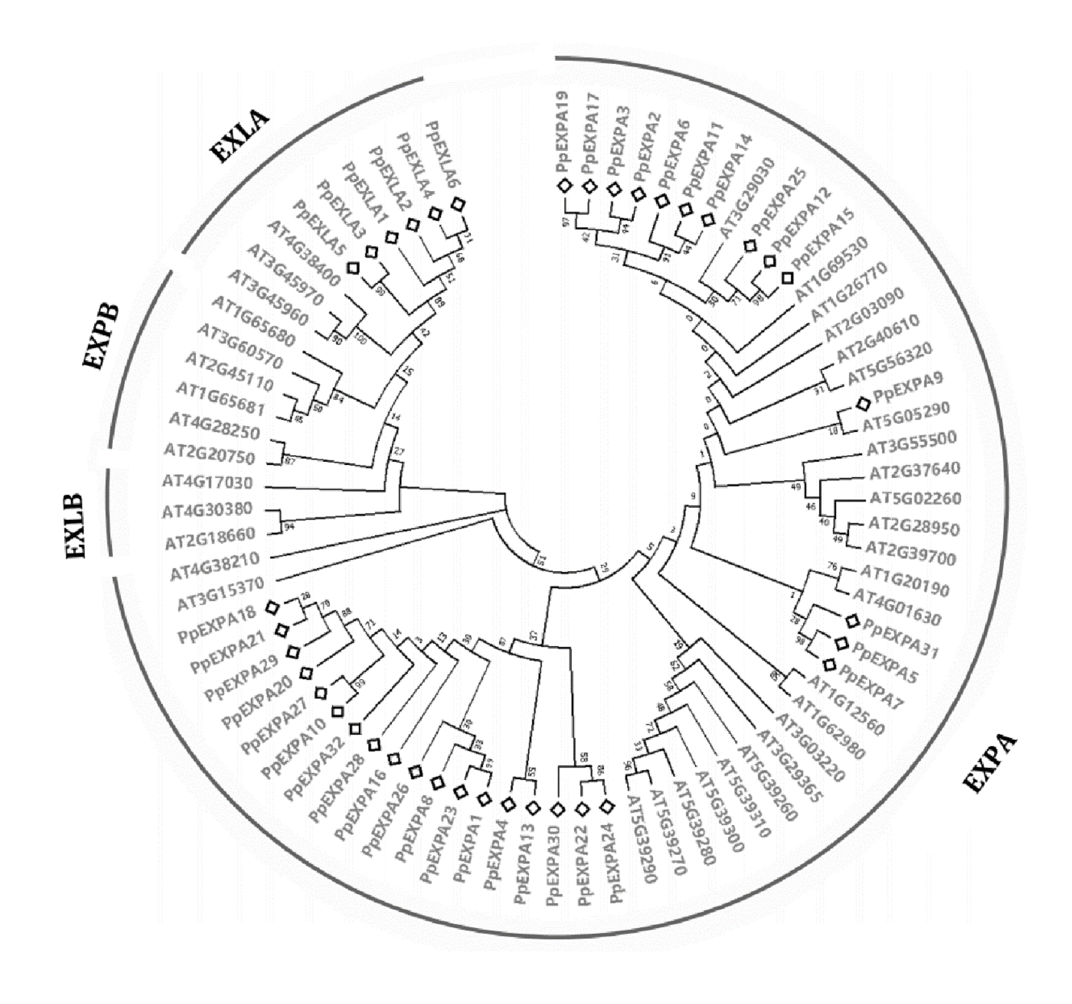


图 3 小立碗藓和拟南芥 Expansin 基因系统进化分析

Fig. 3 Phylogenetic tree of Expansin gene in *Physcomitrella patens* and *Arabidopsis thaliana*

2.4 小立碗藓 EXP 基因家族的染色体位置

通过 Phytozome 数据库获得小立碗藓 EXP 基因家族染色体定位信息，将小立碗藓中 38 个扩展蛋白定位在 15 条染色体上，图 4 中显示小立碗藓第 8 条染色体和第 14 条染色体上均定位了 6 个 EXP 基因。根据基因簇的定义可发现小立碗藓 EXP 基因家族在第 14 条染色体上出现了小规模的基因簇(Bai, et al., 2002)。另 7 条染色体上均只含 1 个基因，3 条染色体含有 2 个基因，可见小立碗藓 38 个 Expansin 基因在染色体上分布是不均匀的。

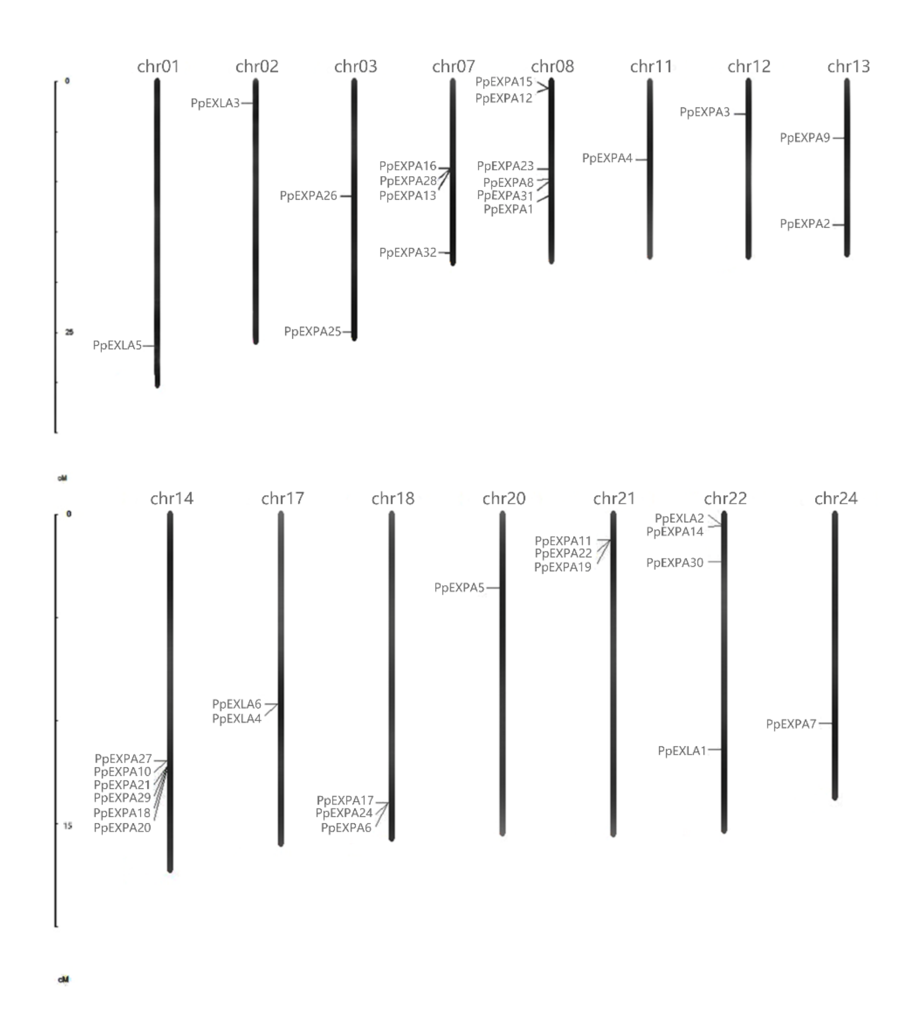


图 4 小立碗藓扩展蛋白基因的染色体定位

Fig. 4 Chromosome distribution of Expansin gene family in *Physcomitrella patens*

3 讨论

扩展蛋白是植物细胞壁的重要组分，主要参与细胞扩张以及一系列发生细胞壁修饰的发育过程以及植物康你的重要生理过程。研究鉴定小立碗藓中含有 38 个扩展蛋白，可进一步分类为 EXPA 亚家族及 EXLA 亚家族。小立碗藓 EXP 基因的两个亚家族之间在基因内含子外显子分布、motif 结构等方面具有一定的保守性，代表了亚家族的重要特征， 研究发现 EXPA 亚家族中约 70%的基因含有 1~ 2 个内含子，另有约 22%的 EXPA 亚家族基因不含内含子；而 EXLA 亚家族几乎仅含 1 个内含子。基序分析表明，EXPA 亚家族及 EXLA 亚家族均含有 motif 2；EXLA 亚家族基因均含有 motif 2 - motif 9 - motif 7 - motif 10 -motif 6 的稳定结构，比 EXPA 亚家族更具有保守性。同一亚家族内的 EXP 基因含有保守的内含子与外显子结构和 motif 特征，而不同亚家族之间 motif 的多样化也表明了小立碗藓 EXP 基因在一定程度上参与细胞内的多种代谢途径。通过染色体定位发现，在小立碗藓 Expansin 基因随机分布 15 条染色体上，除 14 号染色体外，基本不形成基因簇。

EXP 基因在植物体的生长发育及逆境胁迫中具有重要的生理意义，已成为植物体基因功能研究热点之一，但其在小立碗藓中的基因功能尚不明确。本研究通过对已有数据库数据的分析，使用生物学信息的方法手段，筛选得到 38 个小立碗藓扩展蛋白基因家族成员，分析其蛋白特征、进化关系、基因结构及染色体

定位等。从基因水平上展示小立碗藓种属的特征，为后续研究扩展蛋白基因功能提供了理论依据。

参考文献：

BAI J, PENNILL LA, NING JC, et al., 2002. Diversity in nucleotide binding site-leucine-rich repeat genes in cereals [J]. Cytogenet Genome Res, 12(12): 1871-84.

BOLSER DM, RUQURTI B, ROBERTS J, et al., 2017. Ensembl plants: Integrating tools for visualizing, mining, and analyzing plant genomic data [J]. Methods Mol Biol, 1533: 1-31.

BEDINGER P, DURACHKO DM, 1997. Group I allergens of grass pollen as cell wall-loosening agents [J]. Proc Natl Acad Sci USA, 94(12): 6559-6564.

COSGROVE DJ, 2000. Loosening of plant cell walls by expansins [J]. Nature, 407(6802): 321-6. 1997.

COSGROVE DJ, 2005. Growth of the plant cell wall [J]. Nat Rev Mol Cell Biol, 6(11): 850-61.

COSGROVE DJ, LI LC, CHO HT, et al., 2002. The growing world of expansins [J]. Plant Cell Physiol, 43(12): 1436-44.

ERIC JB, BENEDETTO R, JEREMY AR., et al., 2005. Changes in expansin activity and gene expression during ethylene-promoted leaflet abscission in *Sambucus nigra* [J]. J Exp Bot, 56(413): 817-23.

HAN YY, LI AX, LI F, WANG W, 2012. Characterization of a wheat (*Triticum aestivum* L.) expansin gene, TaEXPB23, involved in the abiotic stress response and phytohormone regulation [J]. Plant Physiol Biochem ,54: 49-58.

HANS K, KENT JB, DAVID AB, HYUNG-TAEG C, et al., 2004. Nomenclature for members of the expansin superfamily of genes and proteins [J]. Plant Mol Biol, 55(3): 311-314.

IVICA L, TOBIAS D, PEER B, 2012. SMART 7: Recent updates to the protein domain annotation resource [J]. Nucl Acids Res, 40(Database issue): D302-5.

JAVIER S, ROBERT EC, DANIEL JC, 1992. Two endogenous proteins that induce cell wall extension in plants [J]. The Plant cell, 4(undefined): 1425-33.

JIN HQ, CHEN YH, JIN YF, 2006. The development of expansin (cell wall expansion protein) [J]. Chin Bull Life Sci, 18(2): 168-174. [金慧清，陈英豪，金勇丰，2006. Expansin(细胞壁松弛蛋白)的发展[J]. 生命科学, 18(2): 168-174.]

LI HY, SHI Y, DING YN, et al., 2014. Bioinformatics analysis of expansin gene family in poplar genome[J]. J Beijing For Univ, 36(2):59-67.[李昊阳，施杨，丁亚娜，等, 2014. 杨树扩展蛋白基因家族的生物信息学分析[J]. 北京林业大学学报, 36(2):59-67.]

ROBERT DF, JODY C, SEAN RE, 2011. HMMER web server: interactive sequence similarity searching [J]. Nucl Acids Res, 39(Web Server issue): W29-37.

RENSING SA, ROMBAUTS S, VAN DEP, et al., 2002. Moss transcriptome and beyond [J]. Trends Plant Sci ,7(12): 535-8.

STEFAN AR, DANIEL L, ANDREAS Z, et al., 2008. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants [J]. Science, 319(5859): 64-9.

SAMPEDRO J, CARE YR. E, COSGROVE DJ, 2006. Genome histories clarify evolution of the expansin superfamily: New insights from the poplar genome and pine ESTs [J]. J Plant Res, 119(1): 11-21.

VICTORIA HS, JENNIFER MT, ROBERT EC, 2016. Utility of the *Amborella trichopoda* expansin superfamily in elucidating the history of angiosperm expansins [J]. J Plant Res, 129(2): 199-207.

YI JM, HUANG T, HUANG Y, et al., 2015. Phylogenetic analysis of MADS-box gene family in *Physcomitrella patens*[J]. Plant Physiol J, 51（2）: 197-206.[易吉明, 黄婷, 黄勇, 等, 2015. 小立碗藓 MADS-box 基因家族的系统进化分析[J]. 植物生理学报, 51(2): 197-206.）